



UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE GUAYAQUIL
CARRERA DE COMPUTACIÓN

**IMPACTO DEL BIG DATA EN LA EVALUACIÓN Y OPTIMIZACIÓN DEL
RENDIMIENTO ACADÉMICO**

Trabajo de titulación previo a la obtención del
Título de Ingeniero en ciencias de la computación

AUTOR: MAURICIO ALEJANDRO TURBAY GUSQUI

TUTOR: JOE FRAND LLERENA IZQUIERDO, Ing., Msc.

Guayaquil – Ecuador

2025

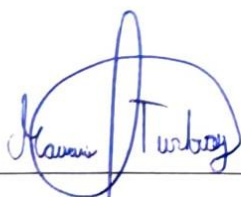
CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Yo, Mauricio Alejandro Turbay Gusqui con documento de identificación N° 0957079403 manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 08 de enero del año 2025

Atentamente,



Mauricio Alejandro Turbay Gusqui

0957079403

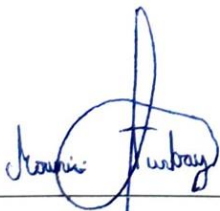
**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Mauricio Alejandro Turbay Gusqui con documento de identificación No. 0957079403, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor(a) del Artículo Académico: “Impacto Del Big Data En La Evaluación y Optimización Del Rendimiento Académico”, el cual ha sido desarrollado para optar por el título de: Ingeniero en Ciencias de la Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 08 de enero del año 2025

Atentamente,



Mauricio Alejandro Turbay Gusqui
0957079403

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Joe Frand Llerena Izquierdo con documento de identificación N° 0914884879, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: IMPACTO DEL BIG DATA EN LA EVALUACIÓN Y OPTIMIZACIÓN DEL RENDIMIENTO ACADÉMICO, realizado por Mauricio Alejandro Turbay Gusqui con documento de identificación N° 0957079403, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 08 de enero del año 2025

Atentamente,



Ing. Joe Frand Llerena Izquierdo
0914884879

DEDICATORIA

Dedico este trabajo a mi familia. Durante ese largo trayecto han sido la fortaleza que he necesito en momento difíciles, no solo a nivel educativo sino en el mundo laboral, me han ayudado a mejorar y lograr mis objetivos personales.

Sin embargo, deseo dar las gracias especialmente a mis padres los cuales han sido los principales actores de que pueda estar redactando este trabajo, me han demostrado y enseñado que el trabajo y la dedicación rompen cualquier obstáculo que se nos presente. Lo primordial en la vida es ser agradecido y mediante esta carta quiero darles mi gratitud por todo lo han hecho y sacrificado por mi desarrollo personal y profesional. Gracias infinitas.

AGRADECIMIENTO

Agradezco profundamente a mi madre, Janeth Amalia Gusqui Parra y a mi padre Mauricio José Turbay Henríquez por la paciencia y dedicación que les han entregado a sus hijos, no es fácil, lo sé ahora que soy adulto, pero tuvieron el coraje, valentía y amor de darme buenos valores y educación que me han ayudado a ser lo que soy ahora. Son mis pilares y lo serán siempre.

RESUMEN

El presente artículo académico analiza la importancia de las soluciones tecnológicas de Big Data en el proceso de evaluación y optimización del rendimiento académico. A continuación, se explica cómo, debido a la digitalización, el paradigma del nuevo sistema educativo debe modificarse para incorporar el uso de plataformas virtuales y otras tecnologías que, a su vez, crean enormes volúmenes de datos. Desafortunadamente, en la mayoría de las situaciones, esta información se pasa por alto, lo que hace que sea imposible descubrir los modelos de enseñanza de los estudiantes y sus fortalezas y debilidades, así como poner en práctica métodos de enseñanza individuales.

La meta del estudio es mostrar el nivel de implementación de Big Data en el sector educativo que se encuentra dentro de la revisión retrospectiva de publicaciones de investigación utilizando el enfoque PRISMA. La atención principal se presta al diseño de estrategias particulares que mejoran el desempeño escolar y la optimización de la formación de los estudiantes. Como resultado aplicado, se presentan soluciones que permiten organizar la información y representarla gráficamente para tomar decisiones pedagógicas bien fundamentadas.

La importancia práctica del estudio es la oportunidad de desarrollar un modelo que pueda ser utilizado en varias instituciones educativas para cambiar radicalmente sus enfoques de enseñanza. Al mejorar no solo el proceso de evaluación de los resultados académicos, sino también crear una red educativa más adaptable y centrada en los requisitos de los estudiantes.

Palabras claves: Big Data, rendimiento académico, transformación digital, estrategias pedagógicas, análisis educativo, sector educativo.

ABSTRACT

This academic article analyses the importance of Big Data technological solutions in the process of evaluating and optimising academic performance. It then explains how, due to digitalisation, the paradigm of the new educational system must be modified to incorporate the use of virtual platforms and other technologies which, in turn, create huge volumes of data. Unfortunately, in most situations, this information is overlooked, making it impossible to discover students' teaching models and their strengths and weaknesses, as well as to put individual teaching methods into practice.

The goal of the study is to show the level of implementation of Big Data in the educational sector found within the retrospective review of research publications using the PRISMA approach. The main attention is paid to the design of particular strategies that improve school performance and the optimisation of students' training. As an applied result, solutions are presented that allow information to be organised and represented graphically in order to make well-founded pedagogical decisions.

The practical significance of the study is the opportunity to develop a model that can be used in various educational institutions to radically change their teaching approaches, improving not only the process of assessing academic results, but also creating a more adaptable educational network focused on student requirements.

Key words: Big Data, academic performance, digital transformation, pedagogical strategies, educational analysis, education sector.

ÍNDICE DE CONTENIDO

1.	INTRODUCCIÓN	10
2.	REVISIÓN DE LITERATURA.....	12
2.1.	El Big Data en el marco de la educación	12
2.2.	Herramientas, técnicas y algoritmos usados por el Big Data en educación	12
2.3.	Desarrollos recientes y desafíos futuros del Big Data en la formación educativa	13
2.4.	Implementación de Big Data en instituciones educativas del Ecuador.....	13
3.	METODOLOGÍA	14
3.1.	Métodos y técnicas de recopilación de datos empleadas	15
3.2.	Métodos y técnicas de análisis de datos	16
4.	RESULTADOS.....	17
5.	DISCUSIÓN.....	26
6.	CONCLUSIÓN	27
	REFERENCIAS	29

1. INTRODUCCIÓN

Teniendo en cuenta el paradigma del aprendizaje profundo como parte del advenimiento de la inteligencia artificial específicamente en el aprendizaje automático (Pereira & Wahi, 2019), se espera que el papel de los macrodatos en el contexto educativo desempeñe un rol más determinante en la mejora de la manera en que se mide el desempeño estudiantil, así como el enfoque general del aprendizaje (Pino Gonzalez & Navarrete Bonilla, 2024). El cambio tecnológico constante de la información y la adopción generalizada de plataformas educativas en línea han dado como resultado la acumulación de enormes cantidades de datos, que van desde el comportamiento de los estudiantes y los resultados obtenidos hasta las tareas realizadas y la retroalimentación (Villao González & López Zambrano, 2024). Sin embargo, un gran número de instituciones educativas no utilizan eficazmente esta información para refinar y mejorar sus técnicas de enseñanza y optimizar la experiencia de aprendizaje (López-Chila, Llerena-Izquierdo, et al., 2023; Salas-Pilco & Yang, 2022).

La evaluación tradicional del aprendizaje se ha caracterizado por su enfoque segmentado, donde "la buena enseñanza y el buen aprendizaje no son necesariamente ideales" y donde el aprendizaje de los alumnos se limita a "lo que se enseña y lo que se aprende" (Yang & Ogata, 2022). Este modelo creó una desconexión entre el proceso de instrucción y los resultados de aprendizaje que deben ser tratados a través de la creación de comunidades de enseñanza y aprendizaje (Ferguson, 2012; López-Chila, Arteaga-Sotomayor, et al., 2023). Esta estrategia promueve una evaluación más orientada al desarrollo de capacidades, que supere el enfoque de "sólo examen, sólo puntuación" y que fomente una mejora continua en los métodos de enseñanza y evaluación (Flanagan et al., 2022; Yang & Ogata, 2022).

La transferencia de conocimientos o el cambio de este es una de las mayores dificultades para que se recopilen y utilicen grandes cantidades de datos de tal manera que favorezca a una mayor comprensión que, a su vez, pueden ser fundamentales para implementar mejores prácticas educativas (Santamaria Chico, 2024). Si bien el uso del estudio de información y el aprendizaje automático ha progresado en varios sectores, el campo de la educación ha sido lento en adoptarlo de manera similar (Musso et al., 2020). Factores como la disponibilidad limitada de infraestructura tecnológica relevante, capacitación adecuada, medidas para minimizar la exposición a mecanismos de restricción de datos y otros han obligado a que el campo quede

rezagado con respecto a otros (Llerena-Izquierdo et al., 2023; Zapata Martínez & Llerena-Izquierdo, 2023).

El objetivo de este trabajo es abordar los desafíos descritos, centrándose en los conocimientos del análisis de Big Data y su relevancia para mejorar el rendimiento académico. Para ello, a partir de la metodología PRISMA, se expondrán acciones para utilizar el Big Data para un análisis del desempeño estudiantil y la evaluación de los resultados logrados mediante intervenciones concretas. La definición del objetivo final busca definir una guía que otras entidades educativas puedan emplear para impulsar el progreso de los sistemas y las mejores prácticas pedagógicas hacia un sistema educativo más adaptable, receptivo y enfocado en el estudiante (P. Wang et al., 2022).

Estos factores constituyen el fundamento continuo que justifica la realización de esta investigación: abordar el problema crítico del bajo rendimiento de los estudiantes, del cual una causa importante es falta de uso de datos disponibles en grandes cantidades y en contextos educativos para evaluar a los estudiantes con el fin de apoyar y proponer nuevas formas de enseñanza, el aprendizaje autónomo y la consecución de resultados para estudios superiores (Peng et al., 2018; Zhao, 2019).

2. REVISIÓN DE LITERATURA

2.1. El Big Data en el marco de la educación

El Big Data ha transformado las percepciones y la gestión del aprendizaje en las instituciones académicas mediante la introducción de nuevas herramientas y metodologías avanzadas para un diagnóstico de grandes cantidades de datos sobre las formas de mejorar la enseñanza y los resultados académicos (Hao, 2021; Llerena-Izquierdo et al., 2024; Peng et al., 2018; Waheed et al., 2020a).

La definición de Big Data dentro del contexto educativo, más específicamente, constituye el análisis masivo de datos provenientes de entornos de aprendizaje e investigación. Según Bellaj et al. (2024), “El término análisis de datos educativos en gran medida se define como el proceso por el cual se recopilan, agregan y se examinan datos de conjuntos masivos”. Estos ayudan a descubrir los patrones de interacciones entre los estudiantes al extraer la información de los archivos de registro excesivamente grandes generados por la mayoría de plataformas y herramientas digitales dirigidas a la educación, con el fin de adecuar las diversas prácticas didácticas con las urgencias propias de los estudiantes.

En este contexto, por lo tanto, el Big Data es no solo el almacenamiento, sino la creación de herramientas analíticas específicas para este campo, así como el desarrollo metodológico eficiente. Munshi & Alhindi (2021) resumen que “Una plataforma integral creada para realizar aplicaciones analíticas en el ámbito del Big Data educativo” es una manera para utilizar los datos como una forma práctica utilizable durante el proceso educacional de los estudiantes. De la misma manera, es importante tener una infraestructura y recursos adecuados para el procesamiento correcto de la información (Imran & Almusharraf, 2024).

2.2. Herramientas, técnicas y algoritmos usados por el Big Data en educación

Apache Hadoop es una de las plataformas que apoya el almacenamiento y procesamiento de grandes datos en la educación a través de su Distributed File System. Belov & Nikulchev (2021) describen que tanto “Hadoop y HDFS habilitan la construcción de arquitecturas como data lakes que maximizan la disponibilidad y accesibilidad de datos para aplicaciones analíticas”. Sin estas soluciones escalables, sería imposible manejar la creciente cantidad de datos generada por las plataformas de aprendizaje y los sistemas académicos (Rane et al., 2024; Saeidnia, 2023).

De la misma manera, poniendo como ejemplo, unas de las herramientas de análisis visual, como los son Tableau y Power BI, las cuales facilitan la conversión de los datos en gráficos interactivos, que los profesores y directores pueden aprovechar para tomar decisiones en favor de una correcta metodología educativa, y que permita aportar de manera positiva a la evolución y adaptación del Big Data en educación (de la Torre & Miguel, 2020; Islam & Ahmed, 2024; Jita & Dhliwayo, 2024; Kulish & Cheng, 2023).

2.3. Desarrollos recientes y desafíos futuros del Big Data en la formación educativa

Hasta ahora, con el avance de la era de la Big Data, la educación ha dado un gran paso con diferentes programas, tal como Google BigQuery (Berisha et al., 2022). No obstante, la aplicación de tales tecnologías se encuentra con una serie de dificultades (Aftab & Siddiqui, 2019; Martins et al., 2020; Ramírez et al., 2019). La deficiencia de infraestructura tecnológica y las deficiencias en la capacitación docente obstaculizan la oportunidad de adoptar herramientas avanzadas (Cui, 2022).

Con tecnologías como la inteligencia artificial, se espera que con el tiempo las interfaces desarrolladas de Big Data permitan cambiar el continuismo de la enseñanza individualizada y pueda evolucionar de tal manera que podamos considerar las particularidades de cada alumno (Z. Han et al., 2020; Hao, 2021; Quadir et al., 2022). Para esto, deberá existir una colaboración entre los actores relevantes para que las barreras actuales sean eliminadas y el potencial real de las metodologías se exponga (C. Wang & Pan, 2024; Zadeh et al., 2021).

2.4. Implementación de Big Data en instituciones educativas del Ecuador

A pesar de que el uso del Big Data en Ecuador en el contexto educativo aún es escaso, su aplicación ha demostrado que permite mejorar el análisis del desempeño académico y optimizar los procesos de enseñanza (Mago & Khan, 2021). Se presenta a continuación un ejemplo concreto de su implementación en una institución educativa del país con la utilización de diferentes modelos y obteniendo resultados satisfactorios.

Entre los casos más significativos destaca el de la Unidad Educativa “Isabel la Católica” en Píllaro, lográndose implementar un sistema de análisis predictivo en base a los datos de uno de 2000 estudiantes entre los años 2014 y 2021. Con el empleo de algoritmos de máquinas de soporte vectorial, árboles aleatorios y regresión logística, se obtuvo una precisión del 96.06 % en la predicción del rendimiento académico. Dicho modelo predijo con precisión a los

estudiantes en riesgo de bajo desempeño estudiantil y facilitó la puesta en marcha de intervenciones de apoyo temprano (Chucuri, 2022).

Además, en otras instituciones del país, el análisis predictivo ha facilitado la optimización de recursos y la personalización del aprendizaje. Mediante herramientas de minería de datos, se han identificado patrones de rendimiento estudiantil, lo cual, ha permitido mejorar la asignación de docentes, actualizar los planes de estudio y diseñar estrategias educativas más efectivas (Chucuri, 2022).

3. METODOLOGÍA

La investigación tiene como objetivo establecer un marco confiable para el estudio del Big Data y su impacto en la evaluación del rendimiento académico. Para ello, se realizará una revisión sistemática de la literatura académica a través de bases educativas confiables e importantes, lo que permitirá un análisis integral de los trabajos seleccionados.

El enfoque examinará el estado actual del uso de Big Data en educación, identificando tanto las estrategias que han tenido éxito como aquellas con presencia en investigaciones previas. Con base en los hallazgos, se propondrán las estrategias para optimizar el rendimiento académico y mejorar el proceso de enseñanza, considerando las necesidades detectadas en los modelos analizados.

Figura 1.

Flujo de revisión sistemática

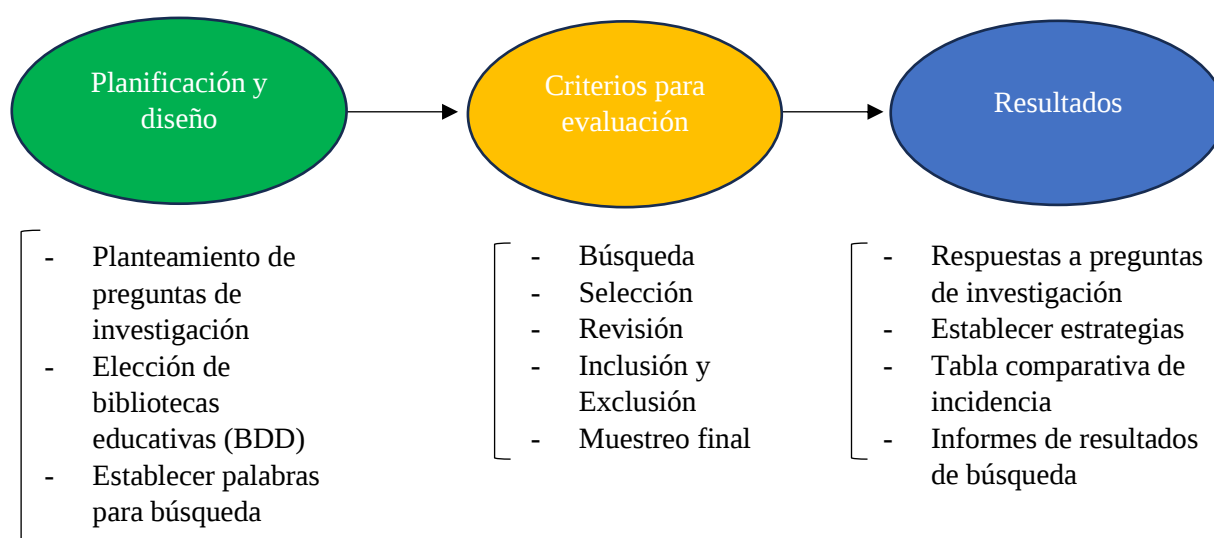


Tabla 1.

Preguntas de investigación

Indicadores	Descripción
P1	¿Qué herramientas, metodologías o algoritmos de Big Data relacionadas con la evaluación del rendimiento académico se utilizan en el sector educativo, y cuál es su diferencial en efectividad?
P2	¿Qué estrategias que implementan patrones de Big Data pueden ser implementadas para afrontar las necesidades identificadas en educación?
P3	¿Cuál es el impacto de las estrategias investigadas sobre el alumnado en el rendimiento y proceso de aprendizaje y, a través de una tabla comparativa de incidencia, cómo se puede demostrar su efectividad definiendo si es efectiva o no?
P4	¿Cómo se clasifican y categorizan las estrategias de Big Data comprendiendo y evaluando su efectividad en el ámbito educativo?
P5	¿Cuáles son los indicadores clave utilizados para medir el impacto de Big Data en el rendimiento académico?

3.1. Métodos y técnicas de recopilación de datos empleadas

Bajo este esquema de trabajo se establecieron los siguientes criterios para una revisión sistemática transparente y eficaz. Para garantizar la relevancia de los resultados, también se aplicaron criterios de inclusión y exclusión:

Mediante el uso de base de datos académicas, como Scopus e IEEE Xplore debido a su amplia cobertura y su importancia para las áreas de investigación relacionadas con la tecnología educativa y el uso de Big Data.

Se emplearán tácticas de búsqueda concretas a través de términos clave vinculados con el tema de investigación, como "Big Data", "educación", "rendimiento" y "estrategias educativas", con la finalidad de cubrir un espectro extenso de investigaciones relevantes.

Para los criterios de inclusión, se tomarán en cuenta investigaciones publicadas entre 2020-2024 sobre Big Data en educación, especificando su uso para valorar y mejorar el rendimiento educacional y las estrategias pedagógicas.

Por otro lado, los criterios de exclusión se definirán con base a trabajos académicos fuera del rango establecido, no relacionados con la educación, sin un análisis basado en Big Data, y

aquellas publicaciones débilmente fundamentadas desde el punto de vista empírico o metodológico.

3.2. Métodos y técnicas de análisis de datos

La elección y búsqueda adecuada de la información es un paso importante en asegurar la eficacia de la metodología, que suponga conducir a hacer un trabajo que facilite la visualización precisa de las distinciones entre las distintas estrategias y metodologías utilizadas. Dicho método se concibe para examinar, registrar y analizar los datos recogidos de forma efectiva. Respecto al procedimiento, el análisis se efectuará a través del enfoque PRISMA, ya que es un marco metodológico preciso y reproducible, lo cual permitirá detallar el proceso de inclusión y exclusión de investigaciones, haciendo uso de los filtros para la eliminación de duplicados, en función de contribuir a un análisis lógico y estructurado.

De este modo, este documento presenta un estudio que contempla una tabla comparativa de la incidencia, obteniendo la correlación de diversas técnicas y estrategias de Big Data en la mejora del desempeño estudiantil. Así podrá observarse los patrones constantes, las tendencias y las áreas que podrían ser optimizadas en las prácticas escolares basadas en datos.

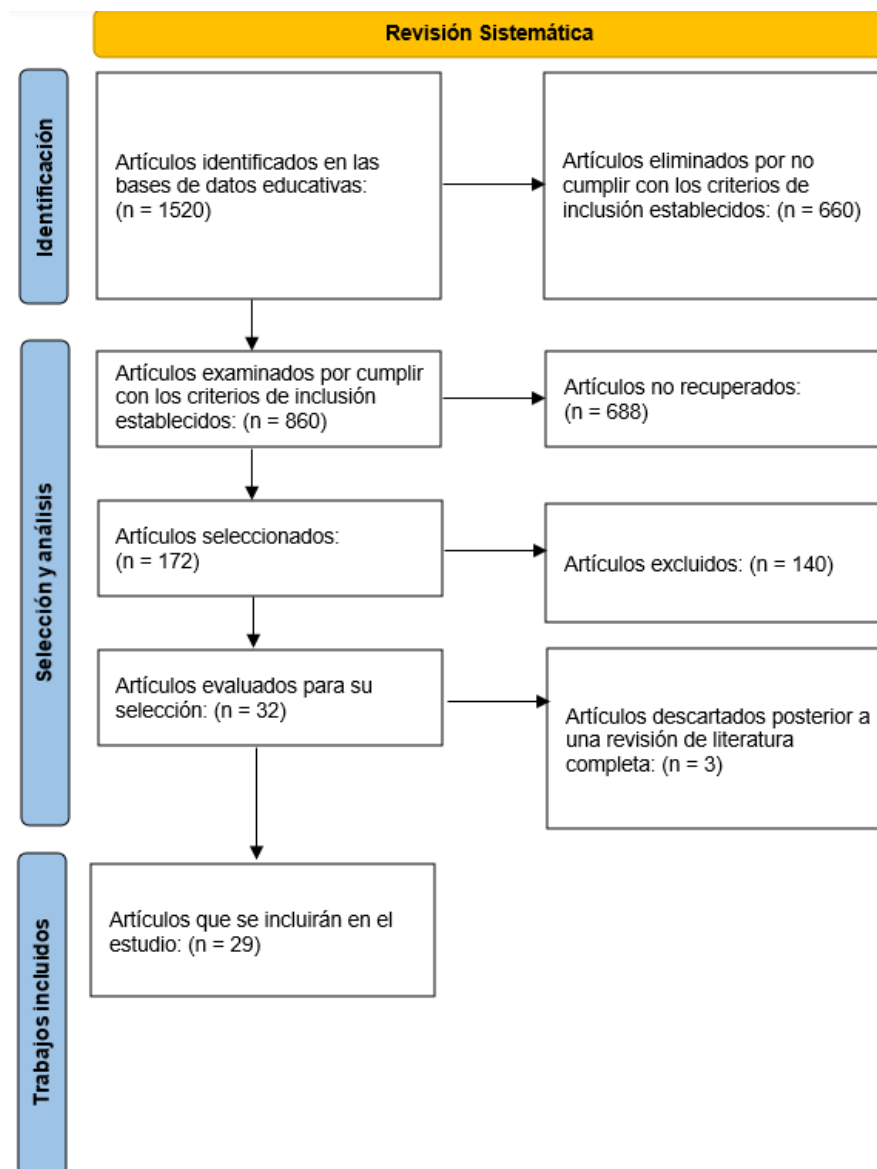
De tal forma, que con base en la evaluación realizada se generará una propuesta de estrategias basadas en Big Data para dar respuesta a las necesidades expuestas. De modo que, los métodos planteados se relacionan directamente con el mejoramiento del rendimiento escolar y el proceso de aprendizaje, donde la clasificación y categorización es un determinante en la generación de soluciones viables.

4. RESULTADOS

Durante la recolección de información, se ha revisado el material disponible en torno al uso de Big Data en la educación, en particular, en lo referente a la mejora del rendimiento estudiantil. Para tener un marco completo, la búsqueda se sustentó en referencias académicas encontradas en las bibliotecas Scopus e IEEE Xplore. Los términos clave de la base fueron “Big Data”, “educación”, “desempeño” y “estrategias educativas”. Obteniendo así un esquema que refleja las diferentes etapas de la revisión sistemática (ver Fig. 2).

Figura 2.

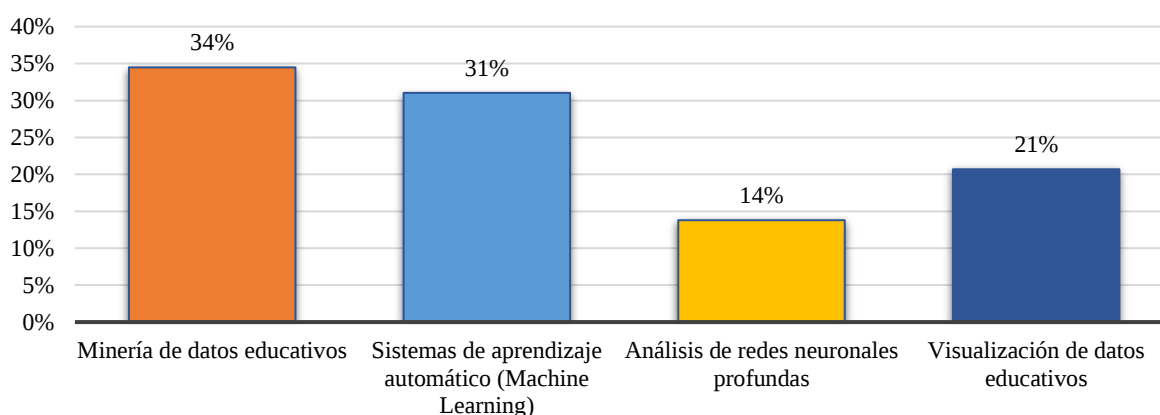
Modelo PRISMA



- a) ¿Qué herramientas, metodologías o algoritmos de Big Data relacionadas con la evaluación del rendimiento académico se utilizan en el sector educativo, y cuál es su diferencial en efectividad?

Figura 3.

P1 - Herramientas y metodologías de Big Data



Minería de datos educativos

El 34% de los documentos inspeccionados destacan la relevancia de esa extracción de datos a la realidad educativa a la luz de los patrones de aprendizaje y las técnicas de influencia sugeridas para mejorar los contextos de aprendizaje. Este planteamiento contribuye de forma notoria a las decisiones informadas en la sectorial educativa (Ozyurt et al., 2023).

Sistemas de aprendizaje automático (Machine Learning)

El 31% de las investigaciones presentadas, respectivamente, mencionan la efectividad de la forma en que el aprendizaje automatizado realiza análisis predictivos y, por ende, identifica a los alumnos en riesgo, mejora en la medida de lo posible el diseño de la pedagogía y provee soluciones adaptativas favorables al aprendizaje personalizado (Rafique et al., 2021).

Análisis de redes neuronales profundas

El 14% de las investigaciones mencionadas sobre el uso de “redes neuronales profundas” se refieren a la aplicación de estas herramientas para resolver problemas complicados en la predicción del desempeño académico. Al mismo tiempo, las herramientas están especialmente adaptadas porque “la minería de datos educativos” implica la recopilación y el análisis de grandes volúmenes de datos no estructurados (Vives et al., 2024).

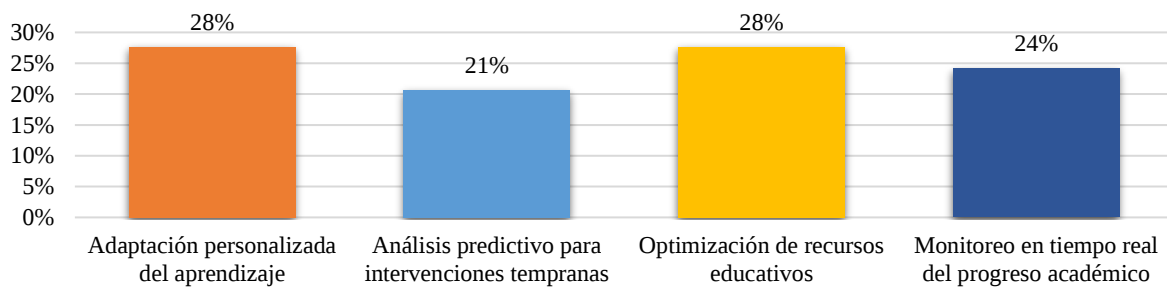
Visualización de datos educativos

El 21% de los textos adoptados presentan la idea de que las capacidades de visualización de datos son importantes para simplificar el aprendizaje en torno a los datos educativos. Llega a ser útil en términos de la posibilidad de que los profesores y puedan comprender rápidamente los resultados de análisis complejos e incluso utilizarlos para tomar decisiones fundamentadas (Feng, Fan, & Ao, 2022).

b) ¿Qué estrategias que implementan patrones de Big Data pueden ser implementadas para afrontar las necesidades identificadas en educación?

Figura 4.

P2 - Estrategias basadas en Big Data



Adaptación personalizada del aprendizaje

El 28% de los textos mencionan la necesidad de personalización para fortalecer la calidad de la experiencia del alumnado. personalizando los contenidos y la experiencia de aprendizaje, las estrategias pueden adaptarse a las necesidades de los alumnos, favoreciendo los estándares de rendimiento (Mafukidze et al., 2024).

Análisis predictivo para intervenciones tempranas

Un 21% de las investigaciones hace un llamado a utilizar el análisis predictivo para detectar aquellos estudiantes con dificultades de riesgo y planificar las intervenciones necesarias con anticipación. las técnicas de predicción basadas en los modelos avanzados de aprendizaje pueden anticipar los comportamientos y disminuir las tasas de abandono educativo (Sridhara et al., 2023).

Optimización de recursos educativos

El 28% de la literatura evaluada explican cómo recursos educativos mejorados eran otro de los beneficios directos que se deriva del uso de Big Data. El objetivo es ayudar a las instituciones educativas a asignar materiales, plataformas y espacios de aprendizaje de manera más eficiente para garantizar un mejor resultado de los estudiantes (Aslam et al., 2021).

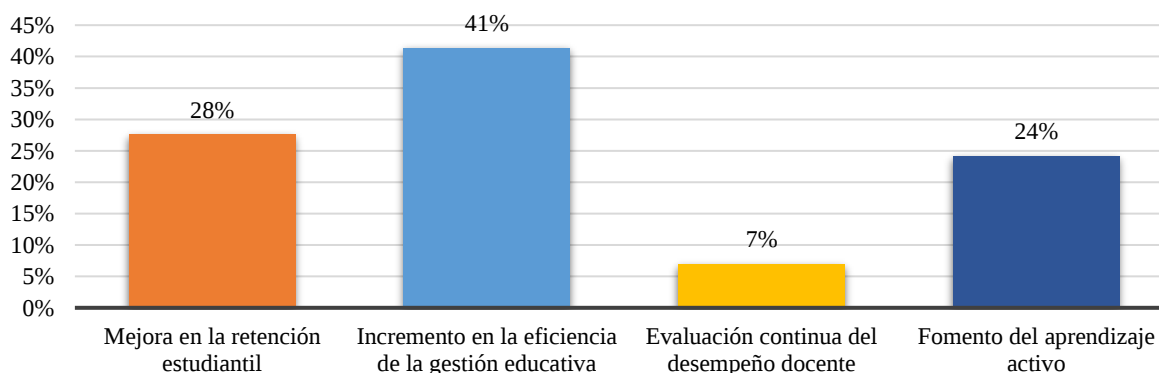
Monitoreo en tiempo real del progreso académico

A su vez, el 24% de los textos subrayan la implementación de sistemas de seguimiento en tiempo real para medir el crecimiento académico. Dichas herramientas ofrecen métricas inmediatamente disponibles para que los profesores puedan ajustar sus estrategias de enseñanza (Rahman et al., 2021).

- c) ¿Cuál es el impacto de las estrategias investigadas sobre el alumnado en el rendimiento y proceso de aprendizaje y, a través de una tabla comparativa de incidencia, cómo se puede demostrar su efectividad definiendo si es efectiva o no?

Figura 5.

P3 - Impacto de las estrategias implementadas basadas en Big Data



Mejora en la retención estudiantil

El 28% de las investigaciones mencionadas enfocan en que el Big Data contribuye significativamente al aumento de la retención y el rendimiento estudiantil estableciendo determinantes de abandono que facilitan las intervenciones oportunas de los entes responsables (Feng, Fan, & Ao, 2022).

Incremento en la eficiencia de la gestión educativa

El 41% de los textos revisados destacan que la eficacia de la administración educativa también es sumamente dependiente del Big Data. Las herramientas de análisis y visualización dejan en claro cuántos, cuándo y dónde deben asignarse recursos, además de apoyar la toma de decisiones para maximización de procesos administrativos y académicos (Ozyurt et al., 2023).

Evaluación continua del desempeño docente

Solamente el 7% de la revisión de literatura atribuye la apreciación continua del trabajo docente como un beneficio directo de la aplicación de Big Data. No obstante, el Big Data tiene el potencial de proporcionar información basada en datos sobre la eficacia de las metodologías de enseñanza. (Vives et al., 2024).

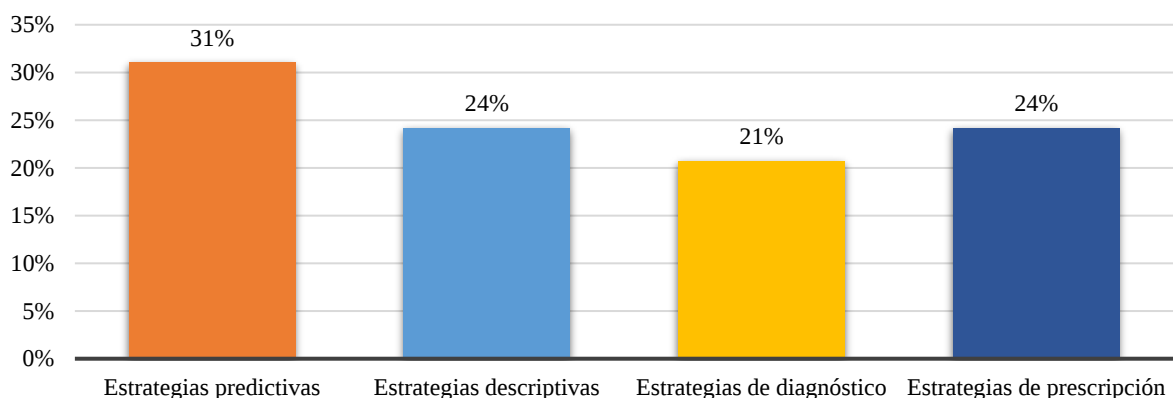
Fomento del aprendizaje activo

El 24% de los artículos destacó que el Big Data apoya el aprendizaje activo y participativo. Significa que la activa participación del estudiantado y la constante interacción significativa con los materiales de enseñanza adaptativos a sus intereses y nivel de habilidades se basa en lo resultante de la utilización del Big Data. (Rafique et al., 2021).

d) ¿Cómo se clasifican y categorizan las estrategias de Big Data comprendiendo y evaluando su efectividad en el ámbito educativo?

Figura 6.

Clasificación y categorización de estrategias de Big Data



Estrategias predictivas

Un 31% de los estudios presentan las estrategias predictivas. Estas son relevantes en el momento de predecir el desempeño académico y tomar las medidas necesarias. Incluyen métodos que utilizan redes neuronales recurrentes y modelos avanzados de estadísticas aplicada para anticipar resultados esenciales para la educación (Sridhara et al., 2023).

Estrategias descriptivas

Un 24% abordan las estrategias descriptivas. Estas buscan comprender los sucesos y patrones comunes del comportamiento de los estudiantes y sus interrelaciones a través de los datos educativos. Con estas estrategias pueden crearse perfiles de aprendizaje para optimizar la experiencia del alumno durante el aprendizaje (Zang, 2024).

Estrategias de diagnóstico

El 21% de los estudios discutidos subraya la necesidad de las estrategias de diagnóstico, que permiten que los métodos identifiquen las brechas en las habilidades y las áreas del aprendizaje que necesitan mejoras. Las estrategias recurren a las herramientas del análisis del clúster y la minería de datos para clasificar a los grupos de estudiantes y las necesidades de cada uno de ellos (Vargianniti & Karpouzis, 2020).

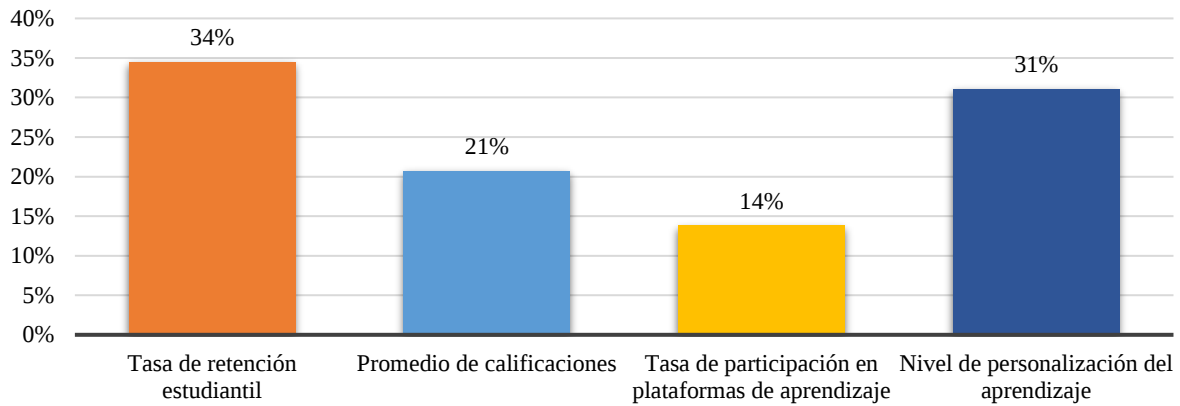
Estrategias de prescripción

Al 24% de los artículos discutidos menciona las estrategias de prescripción. Se utiliza para combinar los métodos predictivos y diagnósticos en cuanto a los métodos recomendados y estudiados. Por lo tanto, les permite a los profesores adaptar los materiales y las actividades curriculares a las necesidades singulares por estudiante (J. Han, 2023).

e) ¿Cuáles son los indicadores clave utilizados para medir el impacto de Big Data en el rendimiento académico?

Figura 7.

Indicadores utilizados para medir/calificar impacto de la Big Data



Tasa de retención estudiantil

El 34% de los trabajos considerados señalan que la tasa de retención de los estudiantes “aumenta significativamente” gracias a las estrategias de esta naturaleza, que permita adelantarse a los desempeños académicos bajos y brindar ayuda oportuna a los estudiantes en riesgo. Dichas estrategias son herramientas analíticas predictivas por medio de algoritmos de clasificación y análisis de tendencias de datos longitudinales (Wu, 2024).

Promedio de calificaciones

El 21% de los trabajos estudiados se refleja que la personalización en el aprendizaje, impulsada por redes neuronales profundas y enfoques de aprendizaje supervisado, influenciada por la personalización en el aprendizaje, eleva el promedio de calificaciones al mejorar la retroalimentación y los contenidos educativos (Fei, 2022).

Tasa de participación en plataformas de aprendizaje

Al 14% de los artículos evaluados presentan que las plataformas educativas tienen un efecto beneficioso en la motivación, así como en el compromiso de estudio de los estudiantes, al tomar el registro de eventos y la minería de datos para rastrear y mejorar la interacción en tiempo real (Altamimi, 2023).

Nivel de personalización del aprendizaje

El 31% de los trabajos consiste en la influencia del grado de personalización en el aprendizaje, que destaca cómo la combinación de técnicas como el aprendizaje supervisado y los análisis de redes de aprendizaje, da la posibilidad de personalizar las experiencias educativas únicas y adecuar el material a las necesidades del estudiantado (Zhang, 2022).

Figura 8.

Análisis PRISMA mediante revisión de literatura

REVISIÓN DE LITERATURA ORIENTADA AL IMPACTO DEL BIG DATA EN LA EVALUACIÓN Y OPTIMIZACIÓN DEL RENDIMIENTO ACADÉMICO																					
Item	Artículos	Herramientas y metodologías de Big Data		Estrategias basadas en Big Data		Impacto de las estrategias implementadas basadas en Big Data		Clasificación y categorización de estrategias de Big Data		Indicadores utilizados para evaluar el impacto de Big Data											
		Mining de datos educativos	Sistemas de aprendizaje automático (Machine Learning)	Análisis de redes neuronales profundas	Visualización de datos educativos	Adaptación personalizada del aprendizaje	Análisis predictivo para intervenciones tempranas	Optimización de recursos educativos	Monitoreo en tiempo real del progreso académico	Medida en la retención estudiantil	Incremento en la eficiencia de la gestión educativa	Evaluación continua del desempeño docente	Fomento del aprendizaje activo	Estrategias predictivas	Estrategias descriptivas	Estrategias de diagnóstico	Estrategias de prescripción	Tasa de retención estudiantil	Promedio de calificaciones	Tasa de participación en plataformas de aprendizaje	Nivel de personalización del aprendizaje
1	2021	Academic Performance Prediction Based on Multisource	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	2022	Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
3	2021	Big Data Platform for Educational Analytics	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0
4	2020	Big Educational Data & Analytics: Survey	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0
5	2021	Deep Learning Model to Predict Students Retention Using BLSM and CBF	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1
6	2024	Design and Optimization of Physical Education Assessment Index System Based on Big Data Analysis	1	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0
7	2024	Development of a Modularized Undergraduate Data Science and Big Data Curricular Using No-Code Software Development Tools	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1
8	2022	Exploration and Visualization of Learning Behavior Patterns from the Perspective of Educational Process Mining	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0
9	2021	Feature Evaluation of Emerging Learning Systems Using Machine Learning: An Extensive Survey	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
10	2021	Impact of Practical Skills on Academic Performance: A Data-Driven Analysis	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0
11	2021	Integrating Learning Analytics and Collaborative Learning for Improving Students' Academic Performance	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1
12	2023	Leveraging Inference: A Regression-Based Learner Performance Prediction System for Knowledge Tracing	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0
13	2024	Prediction of Students' Academic Performance in the Programming Fundamentals Course Using Long-Short-Term Memory Neural Networks	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
14	2023	Uncovering the Educational Data Mining Landscape and Future Perspectives: A Comprehensive Analysis	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
15	2023	A Personalized Learning Path for French Study in Colleges Based on a Big Data Knowledge Map	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1
16	2024	Analysis of Literature Education Strategies and Learning Behaviors in the Age of Big Data	0	1	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0
17	2023	Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0
18	2022	Application of Data Mining Model in Smart Chemistry Education	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0
19	2023	Big Data in Education: Students at Risk as a Case Study	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	0
20	2023	Cognitive Web Service-Based Learning Analytics in Education Systems Using Big Data Analytics	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
21	2022	Datwarehouse design for big data in academia	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
22	2024	Design of Students' Personalized Learning Paths under the Integration and Development of Technology and Basic Education	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
23	2020	Predicting academic performance of students from VLE big data using deep learning models	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0
24	2024	Integration of higher education student management and pedagogical concepts based on data-based decision making	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
25	2022	Personalized Course Recommendation Method Based on Learner Interest Mining in Educational Big Data Environment	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0
26	2020	Using big and open data to generate content for an educational game to increase student performance and interest	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
27	2021	Personalized Education in the Artificial Intelligence Era: What to Expect Next	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
28	2023	Rational Planning of Educational Resources Based on Big Data Fusion	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1
29	2024	The Effectiveness of Educational Resource Optimization Based on Big Data Technology on Enhancing Students' Creative Ability	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0
Sumatoria		10	9	4	6	8	6	8	7	8	12	2	7	9	6	7	10	6	4	9	
Porcentaje		34%	31%	14%	21%	28%	21%	28%	24%	28%	41%	7%	24%	31%	24%	21%	24%	34%	21%	14%	31%
IEEE Xplore																					
Scopus																					

Una vez terminado el análisis de literatura, se presenta una tabla comparativa tomado en base a los trabajos estudiados, donde se detallarán propuestas, métodos y resultados de investigación.

Tabla 2.

Tabla comparativa

Propuesta	Metodología/Algoritmos	Resultados	Referencia
Detección inicial de estudiantes en riesgo en educación autodirigida	Redes Neuronales Artificiales (ANNs)	Efectivo para predecir estudiantes en riesgo	Uliyan et al. (2021), Aslam et al. (2021), Fei (2022)
Predecir el desempeño académico de los estudiantes basado en su salud emocional	Red de Memoria a Largo Plazo (LSTM), Bosque Aleatorio (RF) y Clasificador de Incremento por Gradiente (GBC)	Precisión de predicción del 96%	Vives et al. (2024), Rahman et al. (2021)
Relación entre abandono deportivo, GPA y atribuciones causales	Análisis de perfiles latentes	Rendimiento razonable	Trabelsi et al. (2023), Sridhara et al. (2023)
Predecir el fracaso estudiantil y permitir trayectorias educativas personalizadas	Algoritmos de aprendizaje en conjunto	Mejora en la predicción del fracaso estudiantil	Zhao et al. (2021), Kong (2024)
Detección temprana de estudiantes en riesgo	Clasificador de Vecinos más Cercanos (kNN) y Análisis Discriminante Lineal (LDA)	Buen rendimiento en la detección	Munshi & Alhindi (2021), Vargianniti & Karpouzis (2020)
Predecir el desempeño académico de los estudiantes	RF, vecino más cercano, máquinas de vector soporte (SVM), LR (regresión logística), clasificador Naive Bayes, kNN	Precisión de clasificación entre el 70% y el 75%	Altamimi (2023), Xiao (2023)
Predecir y mejorar el desempeño en MOOCs	Red Bayesiana	Mejora en la predicción del desempeño en MOOCs	Waheed et al. (2020b), J. Han (2023)
Crear un sistema en línea para predecir el desempeño académico	SVM, RF, kNN, ANNs y LR	El desempeño no fue el objetivo principal del artículo. Usaron el error cuadrático medio para medirlo	Jiang (2024), Wu (2024), Dong (2023)
Mejorar la selección de características para predecir el desempeño estudiantil	Algoritmo Genético Binario Mejorado (EBGA)	Mejora en la precisión de predicción	Mafukidze et al. (2024), Rafique et al. (2021)

Predecir el desempeño académico estudiantil usando aprendizaje supervisado	kNN, SVM, Árboles de Decisión (DT), Naïve Bayes (NB), ANNs	Mayor precisión con ANNs (más del 90%)	Zang (2024), Luo (2024)
Predecir el desempeño de estudiantes en educación a distancia	Combinación incremental de clasificadores en conjunto	Mejor rendimiento que clasificadores individuales	Feng, Fan, & Chen (2022), Shijitha et al. (2022)
Predicción temprana de deserción mediante minería de datos	DT, ANNs, SVM	Efectivo en la reducción de tasas de deserción	Ozyurt et al. (2023), Bin (2023), Bhutoria (2022)
Predecir la deserción estudiantil en MOOCs	Modelos temporales	Mejor rendimiento que modelos estáticos	Ang et al. (2020), Zhang (2022)

5. DISCUSIÓN

Desde esta perspectiva, se ha obtenido un análisis de la revisión de la literatura con logros significativos en el manejo del Big Data en la educación. Específicamente, el análisis destaca la adaptación y niveles de personalización del aprendizaje estudiante para una investigación efectiva, la predicción del desempeño académico y el uso eficiente de los recursos educativos. Teniendo como base estas estrategias, podemos decir que proponen resultados o respuestas novedosas a las actuales limitaciones del sistema educativo, mostrando así una amplia capacidad de mejora del desempeño académico y el proceso de enseñanza, reduciendo así un problema como la deserción estudiantil en todos los niveles educativos.

Respecto a la personalización del aprendizaje, es necesario señalar que los patrones predictivos y las redes inteligentes neuronales nos permiten modificar las planificaciones curriculares y las estrategias docentes para adecuarlas a las exigencias individuales del estudiante. Adaptando un método enfocado en la percepción de la información educativa y el apoyo individual llevan a la disminución de los índices de abandono académico (Ang et al., 2020).

Por otra parte, en lo que respecta a la perspectiva del análisis predictivo, diferentes trabajos académicos han resaltado la gran utilidad para los alumnos de estas tecnologías para detectar, de forma anticipada, los factores de riesgo para un rendimiento académico inferior. Este enfoque con un diagnóstico temprano avalaría que es posible implementar rápidamente intervenciones correctivas, incrementando no solo las posibilidades futuras del alumno, sino también la retención escolar y general del desempeño académico (Zhao et al., 2021).

Asimismo, recurrir a mecanismos de inteligencia artificial y procesos algorítmicos de aprendizaje automático para potenciar el uso de recursos educativos permitiría su distribución de forma más eficiente, la mejora de la calidad docente y, finalmente, bajar los costes operativos. Los avances en los algoritmos con inteligencia artificial permiten a las escuelas aprovechar todo su material y dar a los estudiantes solo el que necesiten (Feng, Fan, & Chen, 2022).

6. CONCLUSIÓN

El Big Data utilizado en el ámbito de la educación es una nueva forma de influir en las estrategias y en la productividad académica de los estudiantes reflejada en las calificaciones. En este diagnóstico, se han identificado herramientas y métodos existentes; son los métodos de aprendizaje automático, minería de datos educativos y análisis predictivo, entre otros, es influyente en la evaluación y la optimización del rendimiento académico. El enfoque PRISMA en este estudio ha sido extremadamente útil porque ha ayudado a estructurar y evaluar los artículos correspondientes, permitiendo esclarecer sobre determinadas prácticas relacionadas con la temática de la educación (Liu et al., 2024).

Las instituciones educativas actuales utilizan Big Data de maneras en que adoptan lentamente tecnologías como el aprendizaje profundo, redes neuronales y plataformas de visualización de datos para supervisar y predecir el rendimiento escolar. A pesar de eso, se identificó que todavía hay varias áreas de mejora al respecto de incorporación de tecnologías en beneficio de mejorar y potenciar los métodos pedagógicos del sector estudiantil. La poca atención al entrenamiento adecuado para los docentes y la desconfianza al cambio fue resaltada por la mayoría de las instituciones como una limitante al intentar implementar modelos de visualización de datos (Sun & Song, 2023).

En lo que respecta a las recomendaciones de estrategias basadas en Big Data, la metodología aplicada mediante la revisión de literatura muestra que la personalización de la educación y la utilización de herramientas predictivas son importantes para situar la toma de decisiones de enseñanza a través de los mecanismos adaptativos posibilitados por el análisis de datos resulta en experiencias de aprendizaje más ajustadas al alumno, lo cual impacta en un rendimiento académico superior a la vez que optimiza los materiales y tiempos educativos institucionales (Alkhalil et al., 2021).

Gracias a este enfoque, se logró evaluar con mucha eficacia las distintas estrategias planteadas y ser decisivo para validar el alcance del rendimiento académico. De igual manera, mediante los recursos analíticos como la tabla comparativa de incidencia y los indicadores resultantes del análisis PRISMA, podemos destacar que si se puede precisar la exactitud de la utilidad de las metodologías estudiadas para poder realizar ajustes y mejorar en los métodos de enseñanza. Lo que nos deja, que la implementación que tendría el Big Data en este sector debe ser trabajado con mucha atención para así lograr un éxito en resultado e impacto.

De este modo, se ha logrado un análisis sistemático mediante la revisión de literatura lo que nos ha permitido explorar diferentes estrategias metodológicas que ayuden a potenciar el rendimiento académico de los estudiantes a través de Big Data, los resultados arrojan que los datos en la educación tienen grandes beneficios como la pronta atención en caso de existir un desempeño crítico, sin embargo, todavía existen desafíos que deben ser abordados y desarrollados meticulosamente. Es importante garantizar que la implantación de esta tecnología sea efectiva y se pueda adaptar las necesidades de cada entorno educativo ofreciendo estadísticas que permitan visualizar lo que se puede cambiar o mejorar. La evolución continua de las metodologías y herramientas, del mismo modo la colaboración entre los centros educativos, docentes, investigadores e instituciones gubernamentales es importante para avanzar y conseguir una integración favorable del Big Data en la educación (Rafique et al., 2021).

REFERENCIAS

- Aftab, U., & Siddiqui, G. F. (2019). Big Data Augmentation with Data Warehouse: A Survey. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 2785–2794. <https://doi.org/10.1109/BigData.2018.8622206>
- Alkhalil, A., Abdallah, M. A. E., Alogali, A., & Aljaloud, A. (2021). Applying Big Data Analytics in Higher Education. *International Journal of Information and Communication Technology Education*, 17(3), 29–51. <https://doi.org/10.4018/IJICTE.20210701.oa3>
- Altamimi, A. B. (2023). Big Data in Education: Students at Risk as a Case Study. *Engineering, Technology and Applied Science Research*, 13(5), 11705–11714. <https://doi.org/10.48084/etasr.6190>
- Ang, K. L. M., Ge, F. L., & Seng, K. P. (2020). Big Educational Data Analytics: Survey, Architecture and Challenges. In *IEEE Access* (Vol. 8, pp. 116392–116414). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2020.2994561>
- Aslam, S. M., Jilani, A. K., Sultana, J., & Almutairi, L. (2021). Feature Evaluation of Emerging E-Learning Systems Using Machine Learning: An Extensive Survey. In *IEEE Access* (Vol. 9, pp. 69573–69587). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2021.3077663>
- Bellaj, M., Ben Dahmane, A., Boudra, S., & Lamarti Sefian, M. (2024). Big Data Analytics in Higher Education: A New Adaptive Learning Analytics Model Integrating Traditional Approaches. *International Journal of Interactive Mobile Technologies (IJIM)*, 18(06), 24–39. <https://doi.org/10.3991/ijim.v18i06.46289>
- Belov, V., & Nikulchev, E. (2021). Analysis of Big Data Storage Tools for Data Lakes based on Apache Hadoop Platform. *International Journal of Advanced Computer Science and Applications*, 12(8). <https://doi.org/10.14569/IJACSA.2021.0120864>
- Berisha, B., Mëziu, E., & Shabani, I. (2022). Big data analytics in Cloud computing: an overview. *Journal of Cloud Computing*, 11(1), 24. <https://doi.org/10.1186/s13677-022-00301-w>
- Bhutoria, A. (2022). Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. *Computers and Education: Artificial Intelligence*, 3, 100068. <https://doi.org/10.1016/j.caeai.2022.100068>
- Bin, L. (2023). Cognitive Web Service-Based Learning Analytics in Education Systems Using Big Data Analytics. *International Journal of E-Collaboration*, 19(2). <https://doi.org/10.4018/IJeC.316658>
- Cui, L. (2022). Construction of Big Data Technology Training Environment for Vocational Education Based on Edge Computing Technology. *Wireless Communications and Mobile Computing*, 2022, 1–9. <https://doi.org/10.1155/2022/1060464>
- Chucuri Real, B. (2022). Big Data como herramienta para el análisis del desempeño académico en la educación. [Tesis de Pregrado]. Ambato: Universidad Tecnológica Indoamérica. 83 p. <https://repositorio.uti.edu.ec/handle/123456789/4457>
- de la Torre, E. H., & Miguel, S. G. (2020). Analysis of qualitative data through the system of tables and matrices in educational research. *Revista Electronica Interuniversitaria de Formacion Del Profesorado*, 23(3), 115–132. <https://doi.org/10.6018/REIFOP.435021>

- Dong, Y. (2023). Teaching Quality Monitoring and Evaluation in Higher Education through a Big Data Analysis. *International Journal of Emerging Technologies in Learning (IJET)*, 18(08), 61–78. <https://doi.org/10.3991/ijet.v18i08.39247>
- Fei, S. (2022). Application of Data Mining Model in Smart Chemistry Education. *Wireless Communications and Mobile Computing*, 2022, 1–6. <https://doi.org/10.1155/2022/2506565>
- Feng, G., Fan, M., & Ao, C. (2022). Exploration and Visualization of Learning Behavior Patterns From the Perspective of Educational Process Mining. In *IEEE Access* (Vol. 10, pp. 65271–65283). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2022.3184111>
- Feng, G., Fan, M., & Chen, Y. (2022). Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining. *IEEE Access*, 10, 19558–19571. <https://doi.org/10.1109/ACCESS.2022.3151652>
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5–6), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Flanagan, B., Majumdar, R., & Ogata, H. (2022). Early-warning prediction of student performance and engagement in open book assessment by reading behavior analysis. *International Journal of Educational Technology in Higher Education*, 19(1), 41. <https://doi.org/10.1186/s41239-022-00348-4>
- Han, J. (2023). Rational Planning of Educational Resources Based on Big Data Fusion. *International Journal of Web-Based Learning and Teaching Technologies*, 18(2). <https://doi.org/10.4018/IJWLTT.331086>
- Han, Z., Wu, J., Huang, C., Huang, Q., & Zhao, M. (2020). A review on sentiment discovery and analysis of educational big-data. *WIREs Data Mining and Knowledge Discovery*, 10(1). <https://doi.org/10.1002/widm.1328>
- Hao, J. (2021). Design of Intelligent Educational Administration System Based on Big Data Technology. *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture*, 1390–1394. <https://doi.org/10.1145/3495018.3495406>
- Imran, M., & Almusharraf, N. (2024). Google Gemini as a next generation AI educational tool: a review of emerging educational technology. *Smart Learning Environments*, 11(1), 22. <https://doi.org/10.1186/s40561-024-00310-z>
- Islam, R., & Ahmed, I. (2024). Gemini-the most powerful LLM: Myth or Truth. *2024 5th Information Communication Technologies Conference (ICTC)*, 303–308. <https://doi.org/10.1109/ICTC61510.2024.10602253>
- Jiang, L. (2024). The Effectiveness of Educational Resource Optimization Based on Big Data Technology on Enhancing Students' Creative Ability. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-2792>
- Jita, T., & Dhliwayo, A. (2024). Design-based approach to technology innovation: Teacher educators' experiences with tablets as instructional tools in South Africa. *Contemporary Educational Technology*, 16(3), ep520. <https://doi.org/10.30935/cedtech/14850>

- Kong, J. (2024). Design and Optimization of Physical Education Assessment Index System Based on Big Data Analysis. *IEEE Access*, 12, 73238–73246. <https://doi.org/10.1109/ACCESS.2024.3403139>
- Kulish, O., & Cheng, Y.-Y. (2023). How powerful are Chinese characters' effects on children's creativity? *THINKING SKILLS AND CREATIVITY*, 49. <https://doi.org/10.1016/j.tsc.2023.101374>
- Liu, Y., Tuerdahong, N., & Wang, B. (2024). Research on the Optimization Path of College Education and Teaching Based on Big Data Analysis. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-2862>
- Llerena-Izquierdo, J., Mendez-Reyes, J., Ayala-Carabajo, R., & Andrade-Martinez, C. (2024). Innovations in Introductory Programming Education: The Role of AI with Google Colab and Gemini. In *Education Sciences* (Vol. 14, Issue 12). <https://doi.org/10.3390/educsci14121330>
- Llerena-Izquierdo, J., Rodriguez, M. E., & Guerrero-Roldán, A.-E. (2023). Monitoring and Adaptation of Assessment Activities in a VLE Supported by Learning Analytic. In V. Robles-Bykbaev, J. Mula, & G. Reynoso-Meza (Eds.), *Intelligent Technologies: Design and Applications for Society* (pp. 409–419). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-24327-1_35
- López-Chila, R., Arteaga-Sotomayor, J., Caballero-Barros, E., & Llerena-Izquierdo, J. (2023). An e-learning evaluation method with the effectiveness of the questionnaire tool in the VLEs. *2023 IEEE World Engineering Education Conference (EDUNINE)*, 1–6. <https://doi.org/10.1109/EDUNINE57531.2023.10102811>
- López-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha, N., & Cueva-Estrada, J. (2023). Artificial Intelligence in Higher Education: An Analysis of Existing Bibliometrics. *Education Sciences*, 14(1), 47. <https://doi.org/10.3390/educsci14010047>
- Luo, C. (2024). Analysis of Literature Education Strategies and Learning Behaviors in the Age of Big Data. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1569>
- Mafukidze, H. D., Nechibvute, A., Yahya, A., Badruddin, I. A., Kamangar, S., & Hussien, M. (2024). Development of a Modularized Undergraduate Data Science and Big Data Curricular Using No-Code Software Development Tools. *IEEE Access*, 12, 100939–100956. <https://doi.org/10.1109/ACCESS.2024.3429241>
- Mago, B., & Khan, N. (2021). A Proposed Framework for Big Data Analytics in Higher Education. *International Journal of Advanced Computer Science and Applications*, 12(7). <https://doi.org/10.14569/IJACSA.2021.0120778>
- Martins, A., Martins, P., Caldeira, F., & Sá, F. (2020). *An Evaluation of How Big-Data and Data Warehouses Improve Business Intelligence Decision Making* (pp. 609–619). https://doi.org/10.1007/978-3-030-45688-7_61
- Munshi, A. A., & Alhindi, A. (2021). Big Data Platform for Educational Analytics. *IEEE Access*, 9, 52883–52890. <https://doi.org/10.1109/ACCESS.2021.3070737>
- Musso, M. F., Hernández, C. F. R., & Cascallar, E. C. (2020). Predicting key educational outcomes in academic trajectories: a machine-learning approach. *Higher Education*, 80(5), 875–894. <https://doi.org/10.1007/s10734-020-00520-7>

- Ozyurt, O., Ozyurt, H., & Mishra, D. (2023). Uncovering the Educational Data Mining Landscape and Future Perspective: A Comprehensive Analysis. *IEEE Access*, 11, 120192–120208. <https://doi.org/10.1109/ACCESS.2023.3327624>
- Peng, M. Y.-P., Tuan, S.-H., & Liu, F.-C. (2018). Establishment of Business Intelligence and Big Data Analysis for Higher Education. *Proceedings of the International Conference on Business and Information Management*, 121–125. <https://doi.org/10.1145/3134271.3134296>
- Pereira, A. S., & Wahi, M. M. (2019). Deeper learning methods and modalities in higher education: A 20-year review. *Journal of Higher Education Theory and Practice*, 19(8), 48 – 71. <https://doi.org/10.33423/jhetp.v19i8.2672>
- Pino Gonzalez, E. A., & Navarrete Bonilla, T. M. (2024). *Análisis del uso e impacto de herramientas de inteligencia artificial en el aprendizaje de las ciencias computacionales en la educación superior* [{B.S.} thesis]. <https://dspace.ups.edu.ec/handle/123456789/29276>
- Quadir, B., Chen, N.-S., & Isaías, P. (2022). Analyzing the educational goals, problems and techniques used in educational big data research from 2010 to 2018. *Interactive Learning Environments*, 30(8), 1539–1555. <https://doi.org/10.1080/10494820.2020.1712427>
- Rafique, A., Khan, M. S., Jamal, M. H., Tasadduq, M., Rustam, F., Lee, E., Washington, P. B., & Ashraf, I. (2021). Integrating Learning Analytics and Collaborative Learning for Improving Student's Academic Performance. *IEEE Access*, 9, 167812–167826. <https://doi.org/10.1109/ACCESS.2021.3135309>
- Rahman, M. M., Watanobe, Y., Kiran, R. U., Thang, T. C., & Paik, I. (2021). Impact of Practical Skills on Academic Performance: A Data-Driven Analysis. *IEEE Access*, 9, 139975–139993. <https://doi.org/10.1109/ACCESS.2021.3119145>
- Ramírez, D. M. R., Núñez, S. O. V., Rojas, E. M., & Moreno, H. B. R. (2019). Business intelligence and BigData. *Iberian Conference on Information Systems and Technologies, CISTI, 2019-June*(June), 19–22. <https://doi.org/10.23919/CISTI.2019.8760628>
- Rane, N., Choudhary, S., & Rane, J. (2024). Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation. *Journal of Applied Artificial Intelligence*, 5(1), 69–93. <https://doi.org/10.48185/jaai.v5i1.1052>
- Saeidnia, H. R. (2023). Welcome to the Gemini era: Google DeepMind and the information industry. *Library Hi Tech News, ahead-of-p*(ahead-of-print). <https://doi.org/10.1108/LHTN-12-2023-0214>
- Salas-Pilco, S. Z., & Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: a systematic review. *International Journal of Educational Technology in Higher Education*, 19(1), 21. <https://doi.org/10.1186/s41239-022-00326-w>
- Santamaria Chico, M. A. (2024). *Modelo de un cuadro de mando académico para los profesores basado en las calificaciones de las actividades de los alumnos* [{B.S.} thesis]. <https://dspace.ups.edu.ec/handle/123456789/28724>
- Shijitha, R., Karthigaikumar, P., & Stanly Paul, A. (2022). Data Warehouse Design for Big Data in Academia. *Computers, Materials & Continua*, 71(1), 979–992. <https://doi.org/10.32604/cmc.2022.016676>

- Sridhara, A., Falkner, N., & Atapattu, T. (2023). Leveraging Inference: A Regression-Based Learner Performance Prediction System for Knowledge Tracing. *IEEE Access*, 11, 123458–123475. <https://doi.org/10.1109/ACCESS.2023.3329571>
- Sun, X., & Song, Y. (2023). The Impact of Big Data and AI on Teacher Performance Reviews: A Study of Private Higher Vocational Colleges. *Journal of Information Systems Engineering and Management*, 8(4), 23228. <https://doi.org/10.55267/iadt.07.14050>
- Trabelsi, Z., Alnajjar, F., Parambil, M. M. A., Gochoo, M., & Ali, L. (2023). Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition. *Big Data and Cognitive Computing*, 7(1). <https://doi.org/10.3390/bdcc7010048>
- Uliyan, D., Aljaloud, A. S., Alkhalil, A., Amer, H. S. Al, Mohamed, M. A. E. A., & Alogali, A. F. M. (2021). Deep Learning Model to Predict Students Retention Using BLSTM and CRF. *IEEE Access*, 9, 135550–135558. <https://doi.org/10.1109/ACCESS.2021.3117117>
- Vargianniti, I., & Karpouzis, K. (2020). Using big and open data to generate content for an educational game to increase student performance and interest. *Big Data and Cognitive Computing*, 4(4), 1–20. <https://doi.org/10.3390/bdcc4040030>
- Villao González, J. I., & López Zambrano, E. M. (2024). *Revisión de literatura sobre el uso de Machine Learning enfocados a la seguridad de la información para minimizar vulnerabilidades en el sector educativo* [B.S.] thesis]. <https://dspace.ups.edu.ec/handle/123456789/29275>
- Vives, L., Cabezas, I., Vives, J. C., Reyes, N. G., Aquino, J., Córdor, J. B., & Altamirano, S. F. S. (2024). Prediction of Students' Academic Performance in the Programming Fundamentals Course Using Long Short-Term Memory Neural Networks. *IEEE Access*, 12, 5882–5898. <https://doi.org/10.1109/ACCESS.2024.3350169>
- Waheed, H., Hassan, S.-U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020a). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, 104, 106189. <https://doi.org/10.1016/j.chb.2019.106189>
- Waheed, H., Hassan, S.-U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020b). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, 104, 106189. <https://doi.org/10.1016/j.chb.2019.106189>
- Wang, C., & Pan, J. (2024). Reform and Innovation of College English Teaching Under the Background of Mobile Internet and Big Data. *INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY EDUCATION*, 20(1). <https://doi.org/10.4018/IJICTE.343320>
- Wang, P., Zhao, P., & Li, Y. (2022). Design of Education Information Platform on Education Big Data Visualization. *Wireless Communications and Mobile Computing*, 2022, 1–13. <https://doi.org/10.1155/2022/6779105>
- Wu, J. (2024). Integration of higher education student management and pedagogical concepts based on data-based decision making. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-0991>
- Xiao, G. (2023). A Personalized Learning Path for French Study in Colleges Based on a Big Data Knowledge Map. *Scientific Programming*, 2023. <https://doi.org/10.1155/2023/4359133>

- Yang, C. C. Y., & Ogata, H. (2022). Personalized learning analytics intervention approach for enhancing student learning achievement and behavioral engagement in blended learning. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11291-2>
- Zadeh, A. H., Zolbanin, H. M., & Sharda, R. (2021). Incorporating Big Data Tools for Social Media Analytics in a Business Analytics Course. *Journal of Information Systems Education*, 32(3), 176 – 198. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126940757&partnerID=40&md5=81fb0c58ce32af06e9f1c8f8d8896a04>
- Zang, J. (2024). Design of Students' Personalized Learning Paths under the Integration and Development of Technology and Basic Education. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1737>
- Zapata Martínez, J., & Llerena-Izquierdo, J. (2023). Las TIC después del COVID-19: la perspectiva de los profesores universitarios. *CODES*, 5(SE-Artículos). <https://doi.org/10.15443/codes1988>
- Zhang, R. (2022). Personalized Course Recommendation Method Based on Learner Interest Mining in Educational Big Data Environment. *Scientific Programming*, 2022. <https://doi.org/10.1155/2022/9943965>
- Zhao, L. (2019). Business intelligence application of enhanced learning in big data scenario. *Proceedings - International Joint Conference on Information, Media, and Engineering, IJCIME 2019*, 149–153. <https://doi.org/10.1109/IJCIME49369.2019.00038>
- Zhao, L., Chen, K., Song, J., Zhu, X., Sun, J., Caulfield, B., & Namee, B. Mac. (2021). Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data. *IEEE Access*, 9, 5453–5465. <https://doi.org/10.1109/ACCESS.2020.3002791>